数理解析研究所講究録
1106 巻 1999 年 81-87

Nondeterministic directable automata and related languages

Balázs Imreh* Masami Ito†

The directability of nondeterministic automata can be defined in several nonequivalent ways. In [7], three different notions were introduced and studied. For each case considered, the corresponding directing words constitute a regular language, and thus, the families of such regular languages can be studied. In the work [5], six classes of regular languages are defined in accordance with the different definitions of directability introduced in [7] and the properties of these classes are investigated. In the present work, involving three further classes of regular languages into the investigations, we continue the previous studies and present some further properties with respect to the language families considered. In particular, it is proved that the 8 classes considered constitute a semilattice with respect to the intersection.

1 Introduction

We recall that an input word of an automaton is called *directing* or *synchronizing word* if it brings the automaton from every state into the same state, furthermore, then the automaton considered is called *directable*. The directable automata and directing words have been studied from different points of view (see [2], [3], [5], [6], [7], [8], [10], [12], [13], for example). For nondeterministic (n.d.) automata, the directability can be defined in several ways. We study here three notions of directability which are defined in [7] as follows. An input word w of an n.d. automaton $\mathcal{A}$ is

(1) D1-directing if the set of states aw in which $\mathcal{A}$ may be after reading w consists of the same single state c whatever the initial state a is;

(2) D2-directing if the set aw is independent of the initial state a;

(3) D3-directing if there exists a state c included in all sets aw.

It has to mention that the D_1-directability of complete n.d. automata was already studied by Burkhard [1], where he gave an exact exponential bound for the length of minimum-length D1-directing words of complete n.d. automata. In [5], the classes

*Department of Informatics, József Attila University, Árpád tér 2, H-6720 Szeged, Hungary
†Department of Mathematics, Faculty of Science, Kyoto Sangyo University, Kyoto 603, Japan

of languages consisting of directing words of different types of n.d. automata were studied. Here, we extend our investigations to three further languages and present some new properties of these classes of languages. The paper is organized as follows. The next section provides a general preliminaries, the formal definitions of these languages, and some earlier results. Finally, Section 3 presents some new properties of the language families considered, in particular, it is proved that they constitute a semilattice with respect to the intersection.

2 Preliminaries

Let X be a finite nonempty alphabet. As usual the set of all (finite) words over X is denoted by X^* and the empty word by ε. The length of a word w is denoted by $|w|$.

By an *automaton* we mean a triplet $\mathcal{A} = (A, X, \delta)$, where A is a finite nonempty set of *states*, X is the *input alphabet*, and $\delta : A \times X \longrightarrow A$ is the *transition function.* This function can be extended to $A \times X^*$ in the usual way. By a *recognizer* we mean a system $\mathbf{A} = (A, X, \delta, a_0, F)$, where (A, X, δ) is an automaton, $a_0(\in A)$ is the *initial state*, and $F(\subseteq A)$ is the set of *final states.* The *language recognized* by $\mathbf{A}$ is the set

$$L(\mathbf{A}) = \{w \in X^* : \delta(a_0, w) \in F\}.$$

Sometimes, we say that the recognizer $\mathbf{A}$ *accepts* the language $L(\mathbf{A})$. A language is called *recognizable*, or *regular*, if it is recognized by some recognizer.

An automaton $\mathcal{A} = (A, X, \delta)$ can also be defined as a unary algebra $\mathcal{A} = (A, X)$ for which each input letter x is realized as the unary operation $x^{\mathcal{A}} : A \longrightarrow A$, $a \mapsto \delta(a, x)$. Now, nondeterministic automata can be introduced as generalized automata in which the unary operations are replaced by binary relations. Therefore, by a *nondeterministic (n.d.) automaton* we mean a system $\mathcal{A} = (A, X)$ where A is a finite nonempty set of *states*, X is the *set of the input signs (or letters)*, and each sign $x(\in X)$ is realized as a binary relation $x^{\mathcal{A}}(\subseteq A \times A)$ on A. For any $a \in A$ and $x \in X$, $ax^{\mathcal{A}} = \{b \in A : (a, b) \in x^{\mathcal{A}}\}$ which can be interpreted as the set of states into which $\mathcal{A}$ may enter from state a by reading the input letter x. For any $C \subseteq A$ and $x \in X$, we set $Cx^{\mathcal{A}} = \bigcup\{ax^{\mathcal{A}} : a \in C\}$. This transition can be extended for arbitrary $w \in X^*$ and $C \subseteq A$. This means that if $w = x_1x_2 \ldots x_k$, then $w^{\mathcal{A}} = x_1^{\mathcal{A}}x_2^{\mathcal{A}} \ldots x_k^{\mathcal{A}}(\subseteq A \times A)$. If $w \in X^*$ and $a \in A$, then let $aw^{\mathcal{A}} = \{a\}w^{\mathcal{A}}$.

An n.d. automaton $\mathcal{A} = (A, X)$ is called *complete* if $ax^{\mathcal{A}} \neq \emptyset$, for all $a \in A$ and $x \in X$. Complete n.d. automata are called c.n.d. automata in short. In what follows, we denote a deterministic automaton by $\mathcal{A} = (A, X, \delta)$ and a nondeterministic automaton by $\mathcal{A} = (A, X)$.

The notion of the directability of deterministic automata can be generalized to n.d. automata in several nonequivalent ways. In [7], the following three definitions are presented. Let $\mathcal{A} = (A, X)$ be an n.d. automaton. For any word $w \in X^*$ we consider the following three conditions:

(D1) $(\exists c \in A)(\forall a \in A)(aw^{\mathcal{A}} = \{c\})$;

(D2) $(\forall a, b \in A)(aw^{\mathcal{A}} = bw^{\mathcal{A}})$;

(D3) $(\exists c \in A)(\forall a \in A)(c \in aw^{\mathcal{A}})$.

If w satisfies condition (Di), then w is called a Di-*directing word* of $\mathcal{A}$ ($i = 1, 2, 3$). For every i, $i = 1, 2, 3$, the set of Di-directing words of $\mathcal{A}$ is denoted by $\mathrm{D}_i(\mathcal{A})$, and $\mathcal{A}$ is called Di-*directable* if $\mathrm{D}_i(\mathcal{A}) \neq \emptyset$. It is proved (see [7]) that $\mathrm{D}_i(\mathcal{A})$ is recognizable, for every n.d. automaton $\mathcal{A}$ and i, $i = 1, 2, 3$. The classes of Di-directable n.d. automata and c.n.d. automata are denoted by $\mathbf{Dir}(i)$ and $\mathbf{CDir}(i)$, respectively.

Now, we can define the following classes of languages: For $i = 1, 2, 3$, let

$$\mathcal{L}_{\mathrm{ND(i)}} = \{\mathrm{D}_i(\mathcal{A}) : \mathcal{A} \in \mathbf{Dir}(i)\} \quad \text{and} \quad \mathcal{L}_{\mathrm{CND(i)}} = \{\mathrm{D}_i(\mathcal{A}) : \mathcal{A} \in \mathbf{CDir}(i)\}.$$

Finally, let $\mathbf{D}$ denote the class of directable deterministic automata, and for any $\mathcal{A} \in \mathbf{D}$, let $\mathrm{D}(\mathcal{A})$ be the set of the directing words of $\mathcal{A}$. Moreover, let

$$\mathcal{L}_{\mathrm{D}} = \{\mathrm{D}(\mathcal{A}) : \mathcal{A} \in \mathbf{D}\}.$$

Since all of the languages occuring in the definitions above are recognizable, the defined classes are subclasses of the class of the regular languages.

In what follows, we need the following definition. For any language $L \subseteq X^*$, let us denote by $P_r(L)$ the set of all prefixes of the words in L, *i.e.*, $P_r(L) = \{u : u \in X^* \;\&\; (\exists v \in X^*)(uv \in L)\}$.

Now, we recall some results from [5] and [7] which are used in the following section.

Lemma 1. ([7]) *For any n.d. automaton* $\mathcal{A} = (A, X)$, $\mathrm{D}_2(\mathcal{A})X^* = \mathrm{D}_2(\mathcal{A})$. *If* $\mathcal{A}$ *is complete, then* $X^*\mathrm{D}_1(\mathcal{A}) = \mathrm{D}_1(\mathcal{A})$, $X^*\mathrm{D}_2(\mathcal{A})X^* = \mathrm{D}_2(\mathcal{A})$, *and* $X^*\mathrm{D}_3(\mathcal{A})X^* = \mathrm{D}_3(\mathcal{A})$.

Proposition 1. ([5]) *For a language* $L \subseteq X^*$, $L \in \mathcal{L}_{\mathrm{D}}$ *if and only if* $L \neq \emptyset$, L *is regular, and* $X^*LX^* = L$.

Proposition 2. ([5]) $\mathcal{L}_{\mathrm{CND(2)}} = \mathcal{L}_{\mathrm{D}}$, $\mathcal{L}_{\mathrm{CND(3)}} = \mathcal{L}_{\mathrm{D}}$, $\mathcal{L}_{\mathrm{CND(1)}} \cap \mathcal{L}_{\mathrm{ND(2)}} = \mathcal{L}_{\mathrm{D}}$, and $\mathcal{L}_{\mathrm{CND(1)}} \cap \mathcal{L}_{\mathrm{ND(3)}} = \mathcal{L}_{\mathrm{D}}$.

Furthermore, we need the following proper inclusions which are presented in [5] as well.

Remark 1. ([5]) *The following proper inclusions are valid:*

(a) $\mathcal{L}_{\mathrm{D}} \subset \mathcal{L}_{\mathrm{CND(1)}} \subset \mathcal{L}_{\mathrm{ND(1)}}$,

(b) $\mathcal{L}_{\mathrm{D}} \subset \mathcal{L}_{\mathrm{ND(2)}}$,

(c) $\mathcal{L}_{\mathrm{D}} \subset \mathcal{L}_{\mathrm{ND(3)}}$.

By Proposition 2, $\mathcal{L}_{\mathrm{CND(3)}} = \mathcal{L}_{\mathrm{CND(2)}} = \mathcal{L}_{\mathrm{D}}$, and thus, we investigate the remaining 5 classes and three further classses of languages which are defined as follows. The languages satisfying $X^*L = L$ are called *ultimate definite languages* (*cf.* [9] or [11]), and we shall consider a subclass denoted by $\mathcal{U}$ of this class which consists of all the regular ultimate definite languages. Another additional class, involved into the present investigations, contains all the nonempty regular languages satisfying $P_r(L)LX^* = L$, and this class is denoted by $\mathcal{L}'$. Finally, we also consider the class $\mathcal{L}_{\mathrm{ND(1)}} \cap \mathcal{L}_{\mathrm{ND(3)}}$.

3 Some observations on languages of directing words of n.d. automata

First we consider the classes $\mathcal{U}$ and $\mathcal{L}_{\mathrm{ND}(1)}$. It is known (see [5]) that $\mathcal{L}_{\mathrm{CND}(1)} \subset \mathcal{U}$. On the other hand, $\mathcal{L}_{\mathrm{CND}(1)} \subset \mathcal{L}_{\mathrm{ND}(1)}$ from Remark 1. The following assertion shows that $\mathcal{L}_{\mathrm{CND}(1)}$ is the intersection of these two wider classes.

Proposition 3. $\mathcal{L}_{\mathrm{CND}(1)} = \mathcal{L}_{\mathrm{ND}(1)} \cap \mathcal{U}$.

Using Proposition 2, by the same idea as in the proof of Proposition 3, one can prove the following statement.

Proposition 4. $\mathcal{L}_{\mathrm{ND}(2)} \cap \mathcal{U} = \mathcal{L}_{\mathrm{D}}$ *and* $\mathcal{L}_{\mathrm{ND}(3)} \cap \mathcal{U} = \mathcal{L}_{\mathrm{D}}$.

By the definitions, one can easily obtain the validity of the following observations.

Lemma 2. *If* $L \in \mathcal{L}_{\mathrm{ND}(3)}$, *then* $P_r(L)L = L$ *and* $LP_r(L) = L$.

Lemma 3. *If* $L \in \mathcal{L}_{\mathrm{ND}(1)}$, *then* $P_r(L)L = L$.

As a next step we show that $\mathcal{L}_{\mathrm{ND}(1)}$ and $\mathcal{L}_{\mathrm{ND}(3)}$ are incomparable as well. To verify this statement, let us consider the following examples.

Example 1. Let us define the n.d. automaton $\mathcal{A} = (\{1,2\},\{x,y\})$ by $x^{\mathcal{A}} = \{(1,1),(1,2),(2,1),(2,2)\}$ and $y^{\mathcal{A}} = \{(1,2),(2,2)\}$.

In this case, $\mathcal{A}$ is D_1-directable and $\mathrm{D}_1(\mathcal{A}) = X^*y$. Now, let us suppose that $X^*y \in \mathcal{L}_{\mathrm{ND}(3)}$. Then, $y, xy \in X^*y$ and $x \in P_r(X^*y)$, and thus, Lemma 2 implies that $yx \in X^*y$ which is a contradiction. Therefore, $\mathcal{L}_{\mathrm{ND}(1)} \not\subseteq \mathcal{L}_{\mathrm{ND}(3)}$.

Example 2. Let $\mathcal{A} = (\{1,2\},\{x,y\})$ be the n.d. automaton for which $x^{\mathcal{A}} = \{(1,2),(2,1),(2,2)\}$ and $y^{\mathcal{A}} = \{(1,1)\}$.

Now, $\mathcal{A}$ is D_3-directable and $x, x^2y \in \mathrm{D}_3(\mathcal{A})$ while $xy \notin \mathrm{D}_3(\mathcal{A})$. Let us suppose that $\mathrm{D}_3(\mathcal{A}) \in \mathcal{L}_{\mathrm{ND}(1)}$. Then, there exists an n.d. automaton $\mathcal{B} = (B, X)$ such that $\mathrm{D}_3(\mathcal{A}) = \mathrm{D}_1(\mathcal{B})$. In this case x and x^2y are D1-directing words of $\mathcal{B}$, and thus, there are states $c, d \in B$ such that $Bx^{\mathcal{B}} = \{c\}$, in particular $\{c\}x^{\mathcal{B}} = \{c\}$, and $B(x^2y)^{\mathcal{B}} = \{d\}$. Then, it is easy to see that $B(xy)^{\mathcal{B}} = \{d\}$, and hence, $xy \in \mathrm{D}_1(\mathcal{B}) = \mathrm{D}_3(\mathcal{A})$ must hold which is a contradiction since $xy \notin \mathrm{D}_3(\mathcal{A})$. Consequently, $\mathcal{L}_{\mathrm{ND}(3)} \not\subseteq \mathcal{L}_{\mathrm{ND}(1)}$.

Regarding the new class defined by property $P_r(L)LX^* = L$, the following assertion is valid.

Proposition 5. $\mathcal{L}' = \mathcal{L}_{\mathrm{ND}(2)} \cap \mathcal{L}_{\mathrm{ND}(3)}$.

Proposition 6. $\mathcal{L}' = \mathcal{L}_{\mathrm{ND}(2)} \cap \mathcal{L}_{\mathrm{ND}(1)}$.

From Propositions 5 and 6, the next corollary follows.

Corollary 1. $\mathcal{L}' = (\mathcal{L}_{\mathrm{ND}(1)} \cap \mathcal{L}_{\mathrm{ND}(3)}) \cap \mathcal{L}_{\mathrm{ND}(2)}$.

Since $\mathcal{L}_{\mathrm{ND}(1)}$ and $\mathcal{L}_{\mathrm{ND}(3)}$ are incomparable with respect to the inclusion, $\mathcal{L}_{\mathrm{ND}(1)} \cap \mathcal{L}_{\mathrm{ND}(3)}$ is a proper subclass of both $\mathcal{L}_{\mathrm{ND}(1)}$ and $\mathcal{L}_{\mathrm{ND}(3)}$. On the other hand, by Corollary 1, $\mathcal{L}' \subseteq \mathcal{L}_{\mathrm{ND}(1)} \cap \mathcal{L}_{\mathrm{ND}(3)}$ and $\mathcal{L}' \subseteq \mathcal{L}_{\mathrm{ND}(2)}$. Both of theses inclusions are proper. To verify this observation, let us consider the following examples and arguments.

Example 3. Let the n.d. automaton $\mathcal{A} = (\{1,2\},\{x,y\})$ be defined by $x^{\mathcal{A}} = \{(2,1),(2,2)\}$ and $y^{\mathcal{A}} = \{(1,1),(2,1)\}$.

Then, y is a D1- and D3-directing word, and $y\{y\}^* = \mathrm{D}_1(\mathcal{A}) = \mathrm{D}_3(\mathcal{A})$. Now, if $L \in \mathcal{L}'$, then $P_r(L)LX^*$ must hold which is a contardiction since $y^k x \notin L$, for every integer $k \geq 1$. Therefore, $\mathcal{L}' \subset \mathcal{L}_{\mathrm{ND}(1)} \cap \mathcal{L}_{\mathrm{ND}(3)}$.

Example 4. Let the n.d. automaton $\mathcal{A} = (\{1,2\},\{x,y\})$ be defined by $x^{\mathcal{A}} = \{(1,2),(2,2)\}$ and $y^{\mathcal{A}} = \{(2,1)\}$.
Then, $\mathcal{A}$ is D_2-directable and $\mathrm{D}_2(\mathcal{A}) = xX^* \cup y^2X^*$. Now, if $\mathrm{D}_2(\mathcal{A}) \in \mathcal{L}'$, then since $y \in P_r(\mathrm{D}_2(\mathcal{A}))$ and $x \in \mathrm{D}_2(\mathcal{A})$, $yx \in \mathrm{D}_2(\mathcal{A})$ must hold which is a contradiction. Consequently, $\mathcal{L}' \subset \mathcal{L}_{\mathrm{CND}(2)}$.

By the definition of $\mathcal{L}'$ and Proposition 1, we obviously have that $\mathcal{L}_{\mathrm{D}} \subseteq \mathcal{L}'$. For proving that this inclusion is proper, let us consider the following example.

Example 5. Let $\mathcal{A} = (\{1,2\},\{x,y\})$ be defined by $x^{\mathcal{A}} = \{(2,2)\}$ and $y^{\mathcal{A}} = \{(1,2),(2,2)\}$.

Then, $\mathrm{D}_1(\mathcal{A}) = \mathrm{D}_2(\mathcal{A}) = \mathrm{D}_3(\mathcal{A}) = yX^*$. By Proposition 5, $yX^* \in \mathcal{L}'$. Let us suppose now that $yX^* \in \mathcal{L}_{\mathrm{D}}$. Then, by Proposition 1, $xy \in yX^*$ must hold which a contradiction. Therefore, $yX^* \notin \mathcal{L}_{\mathrm{D}}$, and thus, $\mathcal{L}_{\mathrm{D}} \subset \mathcal{L}'$.

Summarizing, we obtain that taking into account the three new classes, $\mathcal{U}$, $\mathcal{L}'$, and $\mathcal{L}_{\mathrm{ND}(1)} \cap \mathcal{L}_{\mathrm{ND}(3)}$, if $|X| \geq 2$, then we have 8 different classes of regular languages given by Figure 1, where the semilattice of these classes with respect to the intersection is presented.

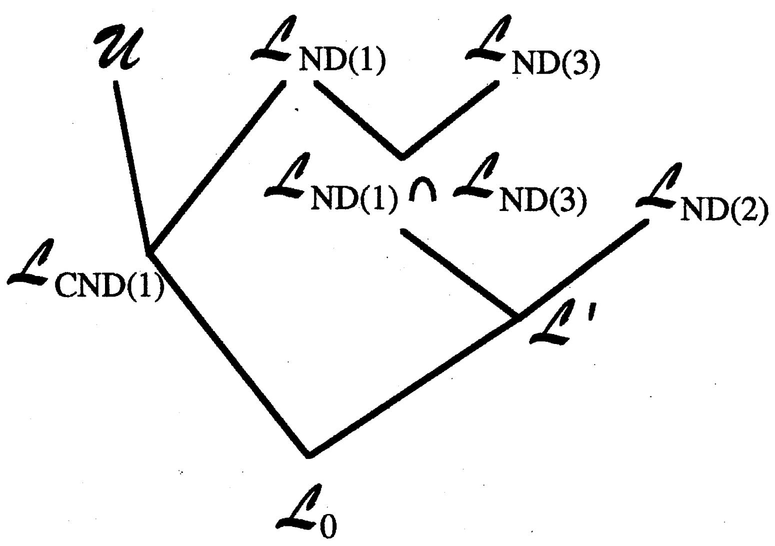

Figure 1

Let $\mathcal{A} = (A, X)$ be a n.d. automaton and $x \in X$. Then, x is called *complete input sign* if $ax^{\mathcal{A}} \neq \emptyset$, for all $a \in A$.

The following statement shows that the languages belonging to $\mathcal{L}_{\mathrm{ND}(2)}$ can be decompose in a particular form.

Proposition 7. *If $L \in \mathcal{L}_{\mathrm{ND}(2)}$, then L is a disjoint union of regular languages L_1 and L_2 where at least one of L_1 and L_2 is nonvoid, furthermore,*

(1) $L_1 \in \mathcal{L}_{\mathrm{D}}$ *provided that* $L_1 \neq \emptyset$

and

(2) $P_r(L_2)L_2Y^* = L_2$ *and* $Y^*L_2Y^* = L_2$ *where* $Y \subseteq X$ *denotes the complete input symbols of* $\mathcal{A}$ *provided that* $Y \neq \emptyset$.

In the rest part of the paper, we study the representation of the languages of $\mathcal{L}_{\mathrm{ND}(2)}$ which have the form $L = MX^*$, where M is a regular prefix code. First we recall some notions.

Let $\emptyset \neq M \subseteq X^+$. Then M is said to be a *prefix code* over X if $M \cap MX^+ = \emptyset$. A prefix code $M \subseteq X^+$ is said to be *maximal* if, for any $u \in X^*$, there exists $v \in X^*$ such that $uv \in MX^*$. Finally, a prefix code M is called *regular* if M is a regular language.

Proposition 8. *Let $M \subseteq X^+$ be a regular prefix code that is not maximal. Let $L = MX^*$. Then $L \in \mathcal{L}_{\mathrm{ND}(2)}$ if and only if $P_r(M)M \subseteq L$.*

The above proposition does not always hold for a regular maximal prefix code.

Example 6. Let $X = \{x, y\}$ and let $A = \{1, 2\}$. Moreover, let $\mathcal{A} = (A, X)$ be the following n.d. automaton: $x^{\mathcal{A}} = \{(1,2), (2,2)\}$, $y^{\mathcal{A}} = \{(1,2)\}$.
Then, $L = D_2(A) = (x \cup yx^*y)X^*y)X^* \in \mathcal{L}_{\mathrm{ND}(2)}$. Let $M = L \backslash LX^*$. Then, $P_r(M)M \subseteq L$ does not hold since $y \in P_r(M)$, $x \in M$ but $yx \notin L = MX^+$.

However, for the class of finite maximal prefix codes, we have the following:

Proposition 9. *Let $\emptyset \neq M \subseteq X^+$ be a finite maximal prefix code. Let $L = MX^*$. Then, $L \in \mathcal{L}_{\mathrm{ND}(2)}$ if and only if $P_r(M)M \subseteq L$.*

Example 7. Let $X = \{x, y\}$ and let $M = \{x, yxx, yxy, yy\}$. Then M is a finite maximal code. Take $y \in P_r(M)$ and $x \in M$. Then, $yx \notin MX^*$. Therefore, $MX^* \notin \mathcal{L}_{\mathrm{ND}(2)}$.

References

[1] H.V. Burkhard, Zum Längenproblem homogener experimente an determinierten und nicht-deterministischen automaten, *Elektronische Informationsverarbeitung und Kybernetik, EIK* **12** (1976), 301-306.

[2] J. Černý, Poznámka k homogénnym experimentom s konečnými automatmi. *Matematicko-fysikalny Časopis SAV* **14** (1964), 208-215.

[3] J. Černý, A. Piricka & B. Rosenauerová, On directable automata, *Kybernetika (Praha)* **7** (1971), 289-297.

[4] S. Eilenberg, *Automata, Languages and Machines*, Academic Press, New York and London, 1974.

[5] B. Imreh, M. Ito, On some special classes of regular languages, to appear.

[6] B. Imreh, M. Steinby, Some remarks on directable automata, *Acta Cybernetica* **12** (1995), 23-35.

[7] B. Imreh, M. Steinby, Directable nondeterministic automata, *Acta Cybernetica*, subbmitted for publication.

[8] M. Ito & J. Duske: On cofinal and definite automata. *Acta Cybernetica* **6** (1983), 181-189.

[9] A. Paz, B. Peleg, Ultimate-definite and symmetric definite events and automata, *J. ACM* **12** (1965), 399-410.

[10] J.-E. Pin, Sur les mots synchronisants dans un automata fini, *Elekronische Informationsverarbeitung und Kybernetik, EIK* **14** (1978), 297-303.

[11] R.G. Reynolds, W.F. Cutlip, Synchronization and general repetitive machines, with applications to ultimate definite automata, *J. ACM* **16** (1969), 226-234.

[12] I. Rystsov, Reset words for commutative and solvable automata, *Theoretical Computer Science* **172** (1997), 273-279.

[13] P.H. Starke, *Abstrakte Automaten*, VEB Deutscher Verlag der Wissenschaften, Berlin 1969.